A Probabilistic Model for Modal Properties based on Operational Modal Analysis

F.L. Zhang[1*], A. M. ASCE; and S.K. Au[2], M. ASCE

Abstract

Operational modal analysis allows one to identify the modal properties (natural frequencies, damping ratios, mode shapes, etc.) of a constructed structure based on output vibration measurements only. For its high economy in implementation it has attracted great attention in theory development and practical applications. In the absence of specific loading information and under uncertain operational environment that can hardly be controlled, the identified modal properties have significantly higher uncertainty than their counterparts based on free or forced vibration tests where the signal-to-noise ratio can be directly controlled. A recent result connecting mathematically the frequentist and Bayesian quantification of identification uncertainty opens up opportunities for modeling the variability of modal properties over time when taking into account identification uncertainty. This paper presents a probabilistic model for the modal properties of a structure under operating environment, which incorporates the identification information from past data to yield the total uncertainty that can be expected in the future with similar structural and environmental characteristics in the past. The developed concepts are illustrated using synthetic, laboratory and field data.

Keywords: uncertainty modeling, operational modal analysis, Bayesian method, ambient modal identification

[1*]Assistant Professor, Research Institute of Structural Engineering and Disaster Reduction, College of Civil Engineering, Tongji University, Shanghai, China (Corresponding author). E-mail: fengliangzhang@hotmail.com, fengliangzhang@tongji.edu.cn; Office phone: +86 21 6598 7352; Office fax: +86 21 6598 2668

[2]Professor, Center for Engineering Dynamics and Institute for Risk and Uncertainty, University of Liverpool, Liverpool, United Kingdom.

Introduction

The 'modal properties' of a structure include primarily its natural frequencies, damping ratios and mode shapes (Clough and Penzien 1993). They govern the structural vibration response under dynamic loads such as wind, earthquake and human excitation. Modern finite element technology has allowed the natural frequencies and mode shapes to be predicted routinely using information available at the design stage. Significant discrepancy from the actual properties can exist, depending on the assumptions made by the engineer. Notably, there is no acceptable method for predicting the damping ratios at the design stage because energy dissipation mechanisms in civil engineering structures are difficult to model using mechanical principles alone. Despite decades of research in structural-wind engineering (Davenport and Hill-Carroll 1986, Jeary 1997, Satake et al. 2003, Kijewski-Correa et al. 2006), there is no accepted method for predicting the damping ratio even for buildings with regular configurations. Blind source separation method is one method recently developed and applied to highly damped structure (Yang and Nagarajaiah 2013) and bridge under traffic loading (Brewick and Smyth 2014). In the past few decades, the damping of tall building has attracted much attention, especially for the amplitude dependence characteristics (Çelebi 1996, Fang et al 1999). Empirical models have been developed to investigate the damping in tall buildings (Bentz and Kijewski-Correa 2008, Spence and Kareem 2014). Damping values used in design are all based on rule of thumb or at best engineering judgment of the design engineer, where uncertainty can arise in this process (Bashor and Kareem 2008). These uncertainties can have significant implications on design economy of modern dynamic-prone structures, which are often featured by creative topology, light-weight materials but met with high performance standards.

Operational modal analysis (OMA) allows one to identify modal properties of a structure using only the ‘output’ vibration (often acceleration) measurements under operating environment. It is recognized as the most economical means for modal identification and has shown promise to be sustainable in civil engineering (Brownjohn 2003, Wenzel and Pichler 2005, Catbas 2011, Reynders 2012). Practical field applications have been reported in many countries and regions. For examples, in the UK, the ambient vibration test of Humber Bridge was carried out by a combined team from different countries and several OMA techniques were used to analyze the data measured (Brownjohn et al. 2010). The results obtained using different methods were investigated and compared. They were also compared with the results of a previous test about 30 years ago and few significant differences were observed in the natural frequencies of the vertical and torsional modes. In the United States, an integral abutment highway bridge was measured by ambient vibration test on the basis of wireless sensors network. OMA with wireless sensor technology was demonstrated for some large civil structures (Whelan et al. 2009). In Portugal, OMA of two historical masonry structures were carried out to estimate the modal parameters and then explore damage assessment (Ramos et al. 2010). In Shanghai, a set of dynamic field tests including ambient and free vibration tests were conducted on a super high-rise building (Shi et al. 2012). By OMA using ambient vibration data, it was found that the damping ratio has a larger discrepancy than natural frequency. The natural frequencies agree well with the finite element model and shaker test result. In Hong Kong, field tests of two super tall buildings were performed under norm and strong wind events (Au et al. 2012). Significant trends were observed between modal parameters and vibration amplitude. The fluctuation of modal parameters induced by environmental effects such as temperature has also attracted increasing attention in the OMA. By one year daily measurement of a tall building, the seasonal variation of modal properties was investigated in Yuen and Kuok (2010). Based on the Timoshenko beam model, a modal frequency-ambient condition model was constructed with the effects of ambient temperature and relative humidity considered. For the Ting Kau Bridge in Hong Kong, one year continuous measurement data have also been collected using accelerometers, temperature sensors, etc. By statistical analysis of the measured frequencies obtained by OMA, it was found that the normal environmental change accounts for variation with variance from 0.2% to 1.52% for the first ten modes of the bridge (Ni et al. 2005). The temperature and humidity effect on the modal parameters of a reinforced concrete slab was also investigated on the basis of OMA results. Clear correlation of natural frequency and damping ratio with temperature and humidity was observed, but it was not obvious for mode shapes (Xia et al. 2006).

The identified modal properties in OMA often have significantly higher uncertainty than their counterparts based on free (zero input) or forced (known input) vibration data because no specific loading information is used in the identification process and the signal-to-noise (s/n) ratio cannot be directly controlled. Quantifying and managing the identification uncertainties then become relevant and important for the proper 'down-stream' use of the modal properties for, e.g., vibration control, structural system identification and more generally structural health monitoring (Papadimitriou et al. 2001, Liu and Duan 2002, Steenackers and Guillaume 2006, Nishio et al. 2012).

OMA has been traditionally performed in a non-Bayesian context. Recent years have seen efforts to quantify and compute the uncertainty of modal properties in a 'frequentist manner' (Pintelon et al. 2007, Reynders et al. 2008, Dohler et al. 2013, El-kafafy et al. 2013), i.e., as the ensemble variability of the identified values over uncertainty of the data. In a more fundamental manner, a Bayesian system identification approach (Beck 2010) allows the information contained in the ambient vibration data to be processed rigorously consistent with modeling assumptions and probability logic to yield inference information on the model parameters of interest. For OMA, a frequency domain approach based on the FFT (Fast Fourier Transform) of ambient vibration data has been recently formulated (Yuen and Katafygiotis 2003) and efficient computational framework has been developed allowing practical applications in a variety of situations, see Au (2011) and Zhang and Au (2013) for well-separated modes, Au (2012a,b) for multiple (possibly close) modes, Au and Zhang (2012a) and Zhang et al. (2015) for well-separated modes in multiple setups. Applications and field studies can be found in Au and To (2012), Au and Zhang (2012b) and Au et al. (2012). A review can be found in Au et al. (2013). By virtual of the one-one correspondence between the time domain data and the FFT, the method allows one to make full use of the relevant information in the data, balancing identification uncertainty and modeling error risk. The results are invariant to the FFTs in other frequency bands which are irrelevant and/or difficult to model.

In applications with a long sequence of ambient vibration data, modal identification is often applied to non-overlapping segments of the data within which the modal properties are assumed to be time-invariant and the stochastic modal force of the modes of interest are assumed to be stationary. A recent mathematical theory connecting the frequentist and Bayesian quantification of uncertainty together with field studies reveals that the posterior uncertainty implied from Bayesian identification of each time segment need not coincide with the ensemble (i.e., segment to segment) variability of the modal properties (Au 2012c). In a Bayesian perspective, the posterior uncertainty obtained is used to quantify the uncertainty of the identified modal parameters, while from a frequentist perspective, the uncertainty is to describe the variability of modal parameters identified among different segments. These two perspectives are two different concepts, and so they need not coincide with each other. The difference is a reflection of modeling error (Au 2012c), which may come from time invariance, damping mechanism, etc., although the source is subject to interpretation and further modeling. With fast Bayesian FFT method, operational modal analysis can be performed to obtain the modal parameters and the associated covariance matrix using ambient vibration data. In this paper we present a probabilistic model that interprets the discrepancy combining Bayesian and frequentist (ensemble concept) perspectives. The new developed model can incorporate the information of the modal parameters and the posterior uncertainty in non-overlapping segments of the data with similar operating environment, which enable the model to assess the distribution of the modal parameters in the future and predict the variation of these parameters. Examples based on synthetic, laboratory and field data are provided to illustrate and apply the developed theory.

Theory

In this section we present a theory for predicting a quantity of interest in a future event using identification information obtained from the past, assuming that the future event is under an uncertain environment that has been experienced by the monitoring database accumulated so far. Although for practical relevance the theory is developed in the context of OMA, it is generally applicable to parameters identified from a Bayesian approach.

Single data set

For instructional purpose, consider first inferring a set of parameters $\boldsymbol{\theta}$ of interest from data D. In a Bayesian context the posterior probability density function (PDF) of $\boldsymbol{\theta}$ acknowledging the information from D and consistent with modeling assumption M and probability logic is given by

$$p(\boldsymbol{\theta} \mid D, M) = P(D)^{-1} P(D \mid \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid M) \tag{1}$$

where, in the conventional terminology, the first term is a normalizing constant, the second term is the likelihood function and the third term is the prior distribution. With sufficient data the prior distribution can be considered slowly varying compared to the likelihood function and so the posterior distribution is directly proportional to the likelihood function. Assuming that the problem is 'globally identifiable', i.e., the posterior distribution has a single peak at the most probable value (MPV) $\hat{\boldsymbol{\theta}}$ (say) in the parameter space of $\boldsymbol{\theta}$. In this case the likelihood function can be approximated by a second order Taylor expansion about the MPV, which gives a Gaussian distribution in the posterior distribution:

$$p(\boldsymbol{\theta} \mid D, M) = (2\pi)^{-n_\theta/2} (\det \hat{\mathbf{C}})^{-1/2} \exp\left[-\frac{1}{2} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \hat{\mathbf{C}}^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \right] \tag{2}$$

where $\hat{\mathbf{C}}$ is the Hessian matrix of the negative of the log-likelihood function (NLLF) evaluated at the MPV; n_θ is the number of parameters in $\boldsymbol{\theta}$. Note that both $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{C}}$ depend on the data D, although this has not been explicitly denoted in the symbol.

For a particular parameter in $\boldsymbol{\theta}$, say, θ, the marginal posterior distribution is also Gaussian:

$$p(\theta \mid D, M) = \frac{1}{\sqrt{2\pi\hat{c}}} \exp\left[-\frac{1}{2\hat{c}} (\theta - \hat{\theta})^2 \right] \tag{3}$$

where $\hat{\theta}$ is the posterior MPV of θ, equal to the corresponding entry in $\hat{\theta}$; and $\hat{c}$ is the posterior variance equal to the corresponding diagonal entry of $\hat{\mathbf{C}}$.

Multiple data sets

Suppose now we have multiple sets of data $D_1,...,D_{N_s}$, where N_s is their number. Although theoretically one can apply the Bayesian method to obtain the posterior PDF $p(\theta | D)$ based on all the data $D = \{D_1,...,D_{N_s}\}$, the result can be misleading because the parameter may have changed over the different time segments and there is a significant chance that the time-invariance assumption implicit in the posterior PDF $p(\theta | D)$ is wrong. That is, the resulting posterior PDF can have significant modeling error that undermines its use.

In view of the above, we relax the invariance assumption and allow the parameter to be different in different data sets. From each data set, say, D_i ($i = 1,...,N_s$), we can perform Bayesian identification in the context of the last section to obtain the posterior PDF of the parameter θ, which is a Gaussian PDF with posterior MPV $\hat{\theta}_i$ and variance $\hat{c}_i$. Although this simple model does not allow us to reach a single posterior PDF to represent the inference information, it has significantly less modeling error than the previous one that assumed invariance. The setting here leads to a 'frequentist' picture of uncertainty, where the different data sets play the role of different realizations of an ensemble population. A simple frequentist measure of the variability of the parameter is in terms of the sample variance of the MPV $\hat{\theta}_i$ over different data sets.

One intuitive question that connects the Bayesian and frequentist perspective of identification uncertainty is whether variability of the MPVs $\{\hat{\theta}_i : i = 1,...,N_s\}$ over different data sets is consistent with the uncertainty implied by the posterior variances $\{\hat{c}_i : i = 1,...,N_s\}$. This question has recently been investigated theoretically, numerically and experimentally (Au 2012c). If there is no modeling error, i.e., the data indeed results from a process following the assumed model and the underlying ('actual') parameters are invariant over the data sets, then the ensemble variance of the MPVs is approximately equal to the ensemble expectation of the posterior variance among the experimental trials. In the context of a finite number of data sets $D_1,...,D_{N_s}$, this means that when N_s is sufficiently large (so that the sample average is close to the ensemble average)

$$\frac{1}{N_s}\sum_{i=1}^{N_s}(\hat{\theta}_i-\bar{\hat{\theta}})^2 \approx \frac{1}{N_s}\sum_{i=1}^{N_s}\hat{c} \tag{4}$$

where

$$\bar{\hat{\theta}} = \frac{1}{N_s}\sum_{i=1}^{N_s}\hat{\theta}_i \tag{5}$$

is the sample average of the MPVs.

In the general case when modeling error can exist due to, e.g., changing experimental conditions or incorrect modeling assumptions, the two quantities need not agree. Conversely, their difference is an indication of modeling error.

Probabilistic modeling of the future

We now develop a probabilistic model for the parameter in a future time window, making use of the information from the data sets $\{D_1,...,D_{N_s}\}$. For this purpose we have to make some 'ergodic' assumptions on the future environment that allow us to relate the future back to the past data, for otherwise a prediction is generally not possible. Specifically, we assume that in a future time window the environment corresponds to either one of the time windows where we have collected the data. Roughly speaking this assumes that the past data is rich enough to cover a future scenario. In the context of OMA, one may have the data for a day, based on which one would like to make a prediction for the modal properties for the next day, which is believed to be under similar environment.

In the above context, we now derive the probability distribution of a generic modal parameter Θ for a future time window. This is denoted by $p_\Theta(\theta \mid D)$, where $D = \{D_1, ..., D_{N_s}\}$ is the collection of all data sets in the monitoring database. This should be distinguished from $p(\theta \mid D)$, which in the context of the previous sections denotes the posterior PDF using all the data sets and assuming that the parameter is invariant over all the time segments (often a poor assumption). The derivation is based on the fact that the future environment is assumed to correspond to one of the time segments observed in the data set with uniform probability of occurrence. Conditional on a given time segment with data D_i (say) the PDF of Θ is simply the posterior PDF $p(\theta \mid D_i)$. Let I denote the index of the time segment that the future event may belong to. It is a random variable uniformly distributed on $\{1, ..., N_s\}$. Using the theorem of total probability,

$$p_\Theta(\theta \mid D) = \sum_{i=1}^{N_s} p_\Theta(\theta \mid I = i, D) P(I = i \mid D) \tag{6}$$

Note that $P(I = i \mid D) = 1/N_s$. Also, when $I = i$, only the data set D_i in D is informative about Θ and so

$$p_\Theta(\theta \mid I = i, D) = p_\Theta(\theta \mid I = i, D_i) = p(\theta \mid D_i) \tag{7}$$

Since $p(\theta \mid D_i)$ is a Gaussian PDF with MPV $\hat{\theta}_i$ and variance $\hat{c}_i$, we have

$$p_\Theta(\theta \mid I = i, D) = \frac{1}{\sqrt{2\pi \hat{c}_i}} \exp\left[-\frac{1}{2\hat{c}_i} (\theta - \hat{\theta}_i)^2 \right] \tag{8}$$

Substituting into equation (6),

$$p_\Theta(\theta \mid D) = \frac{1}{N_s}\sum_{i=1}^{N_s}\frac{1}{\sqrt{2\pi\hat{c}_i}}\exp\left[-\frac{1}{2\hat{c}_i}(\theta-\hat{\theta}_i)^2\right] \tag{9}$$

which is a mixture of Gaussian PDFs.

This model combines the Bayesian and frequentist features of the problem. While the summand in equation (9) is the posterior PDF (Bayesian) based on information in each data segment, the average accounts for the ensemble variability of conditions over different segments. The simple average arises directly from the assumption that the future corresponds to either one of the segments with equal probability, which is justified when e.g., the time segments have equal length and there is no further information suggesting otherwise which time segment is more likely than the others. Note that a mixture distribution of Gaussian PDFs is not Gaussian and it need not be uni-modal.

Bayesian operational modal analysis

In the model developed in the last section, the most probable value and posterior variance of the modal parameters are indispensable. In this section, in the context of OMA we present a fast Bayesian FFT method for determining these quantities. See Au et al. (2013) for an overview, Au (2011) for the well-separated modes and Au (2012a, b) for the general multiple modes. The theory will be briefly outlined as follows.

In the context of Bayesian inference, the measured acceleration is modeled as

$$\ddot{\mathbf{y}}_j = \ddot{\mathbf{x}}_j + \mathbf{e}_j \tag{10}$$

where $\ddot{\mathbf{x}}_j \in R^n$ (j=1, 2,…, N) is the model acceleration response of the structure depending on modal parameters $\boldsymbol{\theta}$ to be identified; and $\mathbf{e}_j \in R^n$ is the prediction error accounting for the discrepancy between the model response and measured data, respectively. The set of parameters $\boldsymbol{\theta}$ includes the natural frequencies, damping ratios, parameters characterizing the power spectral density (PSD) matrix of modal forces, the PSD of prediction error and the mode shape. Here, N is the number of sampling points; n is number of measured degrees of freedom (dofs). The FFT of $\ddot{\mathbf{y}}_j$ is defined as

$$\mathcal{F}_k = \sqrt{\frac{2\Delta t}{N}} \sum_{j=1}^{N} \ddot{\mathbf{y}}_j \exp[-2\pi \mathbf{i} \frac{(k-1)(j-1)}{N}] \tag{11}$$

where $\mathbf{i}^2 = -1$; Δt denotes the sampling interval; $k = 1, ..., N_q$ with $N_q = \mathrm{int}[N/2] + 1$, and $\mathrm{int}[.]$ denotes the integer part; N_q corresponds to the frequency index at the Nyquist frequency.

Let $\mathbf{Z}_k = [\mathrm{Re}\,\mathcal{F}_k; \mathrm{Im}\,\mathcal{F}_k] \in R^{2n}$ be an augmented vector of the real and imaginary part of $\mathcal{F}_k$. In practice, only the FFT data confined to a selected frequency band dominated by the target modes is used for identification. Let such collection be denoted by $\{\mathbf{Z}_k\}$. Using Bayes' Theorem, the posterior PDF of $\boldsymbol{\theta}$ given the data is given by,

$$p(\boldsymbol{\theta} | \{\mathbf{Z}_k\}) \propto p(\boldsymbol{\theta}) p(\{\mathbf{Z}_k\} | \boldsymbol{\theta}) \tag{12}$$

where $p(\boldsymbol{\theta})$ is the prior PDF that reflects the plausibility of $\boldsymbol{\theta}$ in the absence of data. Assuming no prior information, the prior PDF is a constant of $\boldsymbol{\theta}$ and so the posterior PDF $p(\boldsymbol{\theta} | \{\mathbf{Z}_k\})$ is directly proportional to the likelihood function $p(\{\mathbf{Z}_k\} | \boldsymbol{\theta})$. The MPV of modal parameters $\boldsymbol{\theta}$ is the one that maximizes $p(\boldsymbol{\theta} | \{\mathbf{Z}_k\})$ and hence $p(\{\mathbf{Z}_k\} | \boldsymbol{\theta})$.

For large N and small Δt, it can be shown that the FFT at different frequencies are asymptotically independent and their real and imaginary parts follow a Gaussian distribution (Schoukens and Pintelon 1991; Yuen and Katafygiotis 2003). The likelihood function $p(\{\mathbf{Z}_k\}|\boldsymbol{\theta})$ is then a Gaussian PDF of $\mathbf{Z}_k$ with zero mean and covariance matrix $\mathbf{C}_k$ (say). For convenience in analysis and computation, it is written as

$$p(\{\mathbf{Z}_k\}|\boldsymbol{\theta}) \propto \exp[-L(\boldsymbol{\theta})] \tag{13}$$

where $L(\boldsymbol{\theta})$ is the negative log-likelihood function (NLLF):

$$L(\boldsymbol{\theta}) = \frac{1}{2}\sum_k [\ln\det \mathbf{C}_k(\boldsymbol{\theta}) + \mathbf{Z}_k{}^T \mathbf{C}_k(\boldsymbol{\theta})^{-1}\mathbf{Z}_k] \tag{14}$$

The covariance matrix $\mathbf{C}_k$ depends on $\boldsymbol{\theta}$ and is given by

$$\mathbf{C}_k = \frac{1}{2}\begin{bmatrix} \boldsymbol{\Phi}\,\mathrm{Re}\,\mathbf{H}_k\boldsymbol{\Phi}^T & -\boldsymbol{\Phi}\,\mathrm{Im}\,\mathbf{H}_k\boldsymbol{\Phi}^T \\ \boldsymbol{\Phi}\,\mathrm{Im}\,\mathbf{H}_k\boldsymbol{\Phi}^T & \boldsymbol{\Phi}\,\mathrm{Re}\,\mathbf{H}_k\boldsymbol{\Phi}^T \end{bmatrix} + \frac{S_e}{2}\mathbf{I}_{2n} \tag{15}$$

where $\boldsymbol{\Phi} = [\boldsymbol{\Phi}_1,...,\boldsymbol{\Phi}_m] \in R^{n\times m}$ is the mode shapes matrix; m is the number of modes; S_e is the PSD of the prediction error; $\mathbf{I}_{2n}$ denotes the $2n\times 2n$ identity matrix; $\mathbf{H}_k \in C^{m\times m}$ is the theoretical spectral density matrix of the modal acceleration responses and it is given by

$$\mathbf{H}_k = diag(\mathbf{h}_k)\ \mathbf{S}\ diag(\mathbf{h}_k^*) \tag{16}$$

Here $\mathbf{S} \in C^{m\times m}$ denotes the PSD matrix of modal forces; $\mathbf{h}_k \in C^m$ denotes a vector of modal transfer functions with the i-th element equal to

$$h_{ik} = [(\beta_{ik}^2 - 1) + \mathbf{i}(2\zeta_i\beta_{ik})]^{-1} \tag{17}$$

and $diag(\mathbf{h}_k) \in C^{m\times m}$ is a diagonal matrix with the i-th diagonal element equal to h_{ik}; '*'

denotes conjugate transpose; $\beta_{ik} = f_i / \mathrm{f}_k$ denotes frequency ratio; f_i denotes the i-th natural frequency; f_k is the FFT frequency abscissa; ζ_i denotes the i-th damping ratio. The (i, j)-entry of $\mathbf{H}_k$ is given by

$$\mathbf{H}_k(i, j) = S_{ij} h_{ik} h_{jk}^* \qquad (18)$$

where S_{ij} is the (i, j)-entry of $\mathbf{S}$.

The MPV of modal parameters can be theoretically obtained by minimizing the NLLF directly. However, the minimization process is ill-conditioned and the computational time grows drastically with the dimension of the problem, which renders direct solution based on the original formulation impractical in real applications. To solve the computational problems, fast solutions have been developed recently that allow the MPV of modal parameters and associated posterior covariance matrix to be computed typically in several seconds. The basic idea is to reformulate the NLLF in a canonical form and then the singularity with respect to the prediction error PSD S_e can be resolved with the role of the parameters separated, which allows the most probable mode shapes to be determined almost analytically in terms of the remaining parameters. Consequently, the number of modal parameters to be optimized will not increase with the number of measured dofs and only a small set of parameters needs to be optimized. See Au (2011), Zhang and Au (2013) for details on well-separated modes and Au (2012a,b) on general multiple (possibly close) modes. Analytical expressions have also been derived for the Hessian matrix and posterior covariance matrix can be determined as the inverse of the Hessian matrix. This allows the posterior uncertainty to be computed accurately and efficiently without resorting to finite difference.

For completeness the mathematical structure of the problem is briefly illustrated here. First consider the case of well-separated modes, where it is assumed that the selected frequency band contains only one mode. In this case, $\boldsymbol{\theta}$ consists of only one set of natural frequency f, damping ratio ζ, modal force PSD S, PSD of prediction error S_e and mode shape $\boldsymbol{\varphi} \in R^n$. Based on an eigenvector representation of $\mathbf{C}_k$ with one of the basis parallel to $\boldsymbol{\varphi}$, the NLLF can be reformulated as

$$L(\boldsymbol{\theta}) = -nN_f \ln 2 + (n-1)N_f \ln S_e + \sum_k \ln(SD_k + S_e) + S_e^{-1}(d - \boldsymbol{\varphi}^T \mathbf{A}\boldsymbol{\varphi}) \quad (19)$$

where $\boldsymbol{\varphi}$ is assumed to have unit Euclidean norm, i.e., $\|\boldsymbol{\varphi}\| = (\boldsymbol{\varphi}^T \boldsymbol{\varphi})^{1/2} = 1$; N_f is the number of FFT ordinates in the selected frequency band;

$$D_k = [(\beta_k^2 - 1)^2 + (2\zeta\beta_k)^2]^{-1} \quad (20)$$

with $\beta_k = f / \mathrm{f}_\mathrm{k}$ being is a frequency ratio;

$$\mathbf{A} = \sum_k (1 + S_e / SD_k)^{-1} \mathbf{D}_k \quad (21)$$

$$\mathbf{D}_k = \mathrm{Re}\mathcal{F}_k \, \mathrm{Re}\mathcal{F}_k^T + \mathrm{Im}\mathcal{F}_k \, \mathrm{Im}\mathcal{F}_k^T \quad (22)$$

$$d = \sum_k (\mathrm{Re}\mathcal{F}_k^T \, \mathrm{Re}\mathcal{F}_k + \mathrm{Im}\mathcal{F}_k^T \, \mathrm{Im}\mathcal{F}_k) \quad (23)$$

Since the NLLF in equation (19) is a quadratic form in $\boldsymbol{\varphi}$, minimizing it with respect to $\boldsymbol{\varphi}$ under the norm constraint $\|\boldsymbol{\varphi}\| = 1$ gives the MPV of the mode shape, which is simply equal to the eigenvector of matrix $\mathbf{A}$ with the largest eigenvalue. By this way, only four parameters, i.e., $\{f, \zeta, S, S_e\}$, need to be optimized numerically. Consequently, the computational process is significantly shortened with little dependence on the number of measured dofs n.

When there are multiple modes assumed in the selected band, e.g., closely-spaced modes, the MPV of mode shape cannot be determined by solving an eigenvalue problem directly. The problem is more complicated because it is not necessary for the mode shapes (confined to the measured dofs only) to be orthogonal to each other. By representing the mode shape via a set of orthonormal basis and noting that the subspace spanned by such basis does not exceed the number of modes, it is possible to reduce the complexity. In particular, the mode shape matrix $\mathbf{\Phi} \in R^{n \times m}$ in the selected frequency band is represented as

$$\mathbf{\Phi} = \mathbf{B}'\boldsymbol{\alpha} \tag{24}$$

where $\mathbf{B}' \in R^{n \times m'}$ contains a set of (orthonormal) ‘mode shape basis’ spanning the ‘mode shape subspace’ in its columns; $\boldsymbol{\alpha} \in R^{m' \times m}$ contains the coordinates of each mode shape with respect to the mode shape basis in its columns; $m' \le \min(n, m)$ is the dimension of the mode shape subspace. The MPVs of $\mathbf{B}'$ and $\boldsymbol{\alpha}$ need to be determined in the identification process.

Based on equation (24), the NLLF can be expressed as, after a series of mathematical arguments,

$$L(\boldsymbol{\theta}) = \\ -nN_f \ln 2 + (n-m')N_f \ln S_e + S_e^{-1} d + \sum_k \ln|\det \mathbf{E}_k^{'}| - S_e^{-1} \sum_k \mathcal{F}_k^{*} \mathbf{B}^{'} (\mathbf{I}_{m'} - S_e \mathbf{E}_k^{'\,-1}) \mathbf{B}^{'T} \mathcal{F}_k \tag{25}$$

where

$$\mathbf{E}_k^{'} = \boldsymbol{\alpha} \mathbf{H}_k \boldsymbol{\alpha}^T + S_e \mathbf{I}_{m'} \tag{26}$$

is an m'-by-m' Hermitian matrix. On the basis of equation (25), the dimension of matrix computation involved becomes to be m', which is often much smaller than n in applications.

The NLLF depends on the mode shape basis only through the last term in equation (25), which is a quadratic form. The most probable basis minimizes the quadratic form under orthonormal constraints. Although this does not lead to a standard eigenvalue problem, procedures have been developed that allow the most probable basis to be determined efficiently by Newton iteration. A strategy has been developed for determining the MPV of different groups of parameters, iterating until convergence (Au 2012a).

Illustrative examples

In this section we illustrate the developed concepts using examples based on synthetic, laboratory and field data. The example with synthetic data allows us to investigate the theoretical case when there is no modeling error, i.e., the data indeed results from the assumed mechanism; over the different data sets the modal properties are invariant and the stochastic loading are stationary. The example with laboratory experimental data investigates a similar situation under reasonably controlled environment (up to our knowledge). The example with ambient data applies the theory to the real situation where the environment can hardly be controlled.

SDOF structure (synthetic data)

Consider the horizontal vibration of a SDOF structure. It is assumed to have a stiffness of 39.478 KN/mm, a floor mass of 1000 tons and a damping ratio of 1%. The fundamental natural frequency is calculated to be 1 Hz. The structure is subjected to horizontal excitation modeled by independent and identically distributed (i.i.d.) Gaussian white noise with a one-sided spectral density of 1 $(\mu g)^2/Hz$. The acceleration response is calculated at a sampling rate of 20Hz. The measured acceleration is contaminated by measurement noise modeled by Gaussian white noise with a one-sided spectral density of $100\,(\mu g)^2/Hz$. Ambient acceleration data of 600 seconds duration is available. The set of modal parameters to be identified consists of the natural frequency f, damping ratio ζ, PSD of modal force S and PSD of prediction error S_e. The mode shape Φ is trivially equal to 1.

Figure 1 shows the root singular value spectrum of a typical set of 600 sec data. Since there is only one measured dof, the singular value spectrum coincides with the PSD spectrum. The horizontal bar indicates the selected frequency band whose FFT data shall be used for modal identification and the dot indicates the initial guess for the natural frequency. Using this set of data the MPV and posterior covariance matrix of modal parameters can be calculated. To examine the ensemble (frequentist) statistics of the MPVs among statistically identical experimental trials, we generate 100 i.i.d. sets of data (600 sec each). Correspondingly, 100 'samples' of the MPVs are calculated. Figure 2 shows the identified natural frequencies, damping ratios and modal force PSD corresponding to different setups, where each parameter is shown with a dot at the MPV and an error bar covering +/- 2 posterior standard deviations. The ensemble variability of the identification results of the three modal parameters over different data sets is small.

Table 1 compares the frequentist and Bayesian statistics of the modal identification results among the 100 trials. The second column shows the exact parameter value that generated the data. The third column shows the MPV calculated using a typical data set. The sample mean of the MPVs from 100 data sets is shown in the fourth column. The MPV calculated using a single data set and the sample mean are quite close to the exact value. The fifth column titled 'Freq.' shows the sample coefficient of variation (c.o.v.) of the MPVs among the 100 trials, equal to the sample standard deviation divided by the sample mean of the MPV. The sixth column titled 'Bay.' shows the equivalent mean posterior c.o.v. (defined as the sample root mean square (r.m.s.) value of the posterior standard deviation / the sample mean of the MPV). It can be seen that these two quantities are quite close to each other, with the ratios of the frequentist to Bayesian quantity all close to 1(shown in the column titled 'A/B'). The frequentist result is consistent with the posterior uncertainty of these modal parameters in a Bayesian manner.

Figure 3a) shows the PDF of the modal parameters in a future scenario incorporating the information of the 100 data sets, based on the probabilistic model developed in this work, i.e., equation (9). As mentioned before, a mixture distribution of Gaussian PDFs is not necessarily Gaussian. In the present case, the distribution for the natural frequency and prediction error PSD appears to be approximately Gaussian. The same is not true for the damping ratio or the modal force PSD. The mean and c.o.v. (=standard deviation/mean) of the distribution are shown in the title of each subfigure. The mean and standard derivation (std) of the distribution $p_\Theta(\theta \mid D)$ in equation (9) can be determined in terms of $\{\hat{\theta}_i\}$ and $\{\hat{c}_i\}$ as (see Appendix):

$$\text{mean} = \frac{1}{N_s}\sum_{i=1}^{N_s}\hat{\theta}_i = \text{sample mean of } \{\hat{\theta}_i\}_{i=1}^{N_s} \tag{27}$$

$$\begin{aligned}\text{std} &= \left\{\frac{1}{N_s}\sum_{i=1}^{N_s}\hat{c}_i + \frac{1}{N_s}\sum_{i=1}^{N_s}\hat{\theta}_i^2 - \left(\frac{1}{N_s}\sum_{i=1}^{N_s}\hat{\theta}_i\right)^2\right\}^{1/2} \\ &= \left\{\text{sample mean of } \{\hat{c}_i\}_{i=1}^{N_s} + \text{sample variance of } \{\hat{\theta}_i\}_{i=1}^{N_s}\right\}^{1/2}\end{aligned} \tag{28}$$

Figure 3b) shows that the distributions are similar when 1000 data sets are used. This suggests that the number of data sets (100) is sufficiently large so that the distribution is insensitive to it. On other hand, the shape of the distributions is similar when the duration of data used is doubled (Figure 3c) or halved (Figure 3d). An increase in the data length will lead to a decrease in the c.o.v.s of the modal parameters in the prediction model. This is a reflection of the amount of information used in modal identification.

It is instructive to compare the uncertainty implied by the posterior distribution of a single set of data and that implied by the prediction model in equation (9). Table 2 shows the c.o.v.s calculated based on single data set, where Cases a) to d) correspond to the Cases a) to d) in Figure 3, respectively. It is seen that the c.o.v.s in Figure 3 are larger than those in Table 2. This reveals that the uncertainty implied by the prediction model is higher than the posterior uncertainty based on a single data set, which has not incorporated possible data to data variability.

Laboratory frame

Consider a four-storied shear frame situated in the laboratory, as shown in Figure 4. The frame was kept in an air-conditioned room. Eight uni-axial accelerometers are instrumented at the center of the four floors to measure the response along the weak and strong direction for 20 hours in a quiet environment where there was little human activities nearby. In this work, only the data collected in weak direction are investigated. In the nominal case, the first 10-hour data is divided into 60 sets, each with a time window of 10 minutes. Digital data was originally sampled at 2048 Hz and later decimated to 64 Hz for modal identification.

Figure 5 shows the singular value spectrum of a typical data set and the frequency band selected for modal identification. Bayesian modal identification is performed separately for each data set. Figure 6 shows the identification result of the first mode. The MPV of natural frequency only changes slightly over the 10 hour duration. The ensemble variability of the identification results of the damping ratio and the modal force PSD over different data sets is larger than that of the natural frequency, although the MPVs are still in the same order of magnitude. For the natural frequency and damping ratio, the posterior uncertainty (say, in terms of the length of the error bar) is consistent with the ensemble variability of their MPVs over different setups and therefore the Bayesian and frequentist perspectives roughly agree. However, this is not true for the modal force PSD, as evidenced by the typical observation that there is little overlap between the error bars of neighboring setups. This suggests that the controls applied to the laboratory environment can only maintain the modal force PSD to the same order of magnitude but not the same value (to within identification precision) in each data set.

Table 3 summarizes the frequentist and Bayesian statistics among the 60 data sets. For the natural frequency and damping ratio, the sample c.o.v. may be considered similar to the equivalent mean posterior c.o.v.. This is not the case for the modal force PSD, where the sample c.o.v. is more than twice of the equivalent mean posterior c.o.v.. The sample c.o.v. largely reflects the variability in the environment because the posterior c.o.v. is relatively small, although the environment in the laboratory seems controllable. This result is consistent with Figure 6.

Figure 7a) shows the PDF of the modal parameters based on the proposed probabilistic model (equation (9)) using the 60 sets of identification results, where the mean and c.o.v. of this distribution are shown above each subfigure. The distribution of the natural frequency and damping ratio appears to be roughly Gaussian but the same is hardly true for the PSD of modal force and PSD of prediction error. As a sensitivity study, Figure 7b) to d) show results using the first 10, 30 and 120 data sets. It is found that the distribution of natural frequency and damping ratio are generally similar to those in Figure 7a) but the same is not true for the PSD of modal force and the PSD of prediction error. The sensitivity of the latter quantities is reasonable because according to the proposed model the data incorporated reflects the environment that will be experienced. On the other hand, the distribution of the natural frequency and damping ratios are relatively stable because the environment has not changed significantly to the extent that will affect them.

Next, similar to the synthetic data example, the mean and c.o.v. values of these distributions are also investigated. The posterior c.o.v. of natural frequency, damping ratio, modal force PSD and PSD of prediction error for the first data set are equal to 0.059%, 33.1%, 5.9% and 3.3%, respectively. As before, the posterior c.o.v. for a single data set is lower than that calculated based on the prediction model. In the prediction model, the c.o.v. does not necessarily increase when more data sets are incorporated. For example, the c.o.v. of the damping ratio and PSD of prediction error first increases and then decreases when the number of data set used changes from 10 to 120. This implies that in a similar environment, when the number of data set used is adequate, the prediction model will tend to stabilize.

Super-tall building

Consider a tall building situated in Hong Kong measuring 310 m tall and 50 m by 50 m in plan, as studied in Au and To (2012). Ambient data with a duration of 30 hours are collected using a tri-axial accelerometer (i.e., 3 dofs) under normal wind situation in October 2010. The data is divided into 60 data sets of 30 minutes each. Figure 8 shows the root singular value spectrum of a typical data set and the selected frequency band. Modal identification is performed for each data set separately. In the frequency band indicated in the figure there are two closely spaced modes. These two modes will be identified simultaneously. Only the results of the first mode will be discussed, however.

The identification result is shown in Figure 9, with a dot at the MPV and an error bar of +/- two posterior standard deviations. Compared with the previous two examples, the MPVs of the modal parameters have larger variability over different setups. This clearly demonstrates that the posterior uncertainty in one setup does not necessarily tell what will happen to the identification result in the next setup. There is no guarantee that the MPV of the next setup will lie within the error bars of the current setup. For example, the MPV of the natural frequency in Setup 17 lies beyond the region covered by the error bar of Setup 16. One possibility for this is that the modal properties of the structure have changed from one setup to another. This is especially obvious for the modal force PSD shown in the bottom plot of Figure 9. The short error bars imply that the modal force PSD in each setup can be identified quite accurately (within time-invariant assumption within the time window) but it is changing from one setup to another. Table 4 summarizes the frequentist and Bayesian statistics. The sample c.o.v. of S is much larger than the equivalent mean posterior c.o.v., which is likely attributed to environmental variability over different data sets. Similar to the former two examples, the frequentist and Bayesian statistics are similar for the natural frequencies and damping ratios, although the frequentist statistics is consistently larger.

Figure 10 shows the PDF of the modal parameters based on the proposed probabilistic model (equation (9)) using the 60 sets of identification results, where the mean and c.o.v. of this distribution are shown above each subfigure. It is seen that the distributions of the natural frequency, damping ratio and PSD of prediction error appear to be roughly Gaussian, although the environment in the field is much different from that in the laboratory. The distribution of the modal force PSD is multi-modal. This modal parameter is quite sensitive to the environment. The c.o.v. of this quantity calculated according to the prediction model is about 100%.

Note that in real environment, the probabilistic model has not explicitly taken into account the effect of the environmental conditions directly, e.g., temperature, humidity and so on, although the PSD of modal force and PSD of prediction error involved in the model can also reflect some effect from the environment.

Conclusions

This paper develops a probabilistic model on the basis of the data collected, which combines Bayesian identification results with a frequentist quantification of the setup-to-setup variability. The effect of data length and the number of data sets on the probabilistic model have been investigated. Increasing the data length can reduce posterior uncertainty of modal parameters identified from each data set, to the extent that the stationary assumption within each data set is still valid. Increasing the number of data sets can generally enrich the monitoring data base and provide a more robust prediction of the future. The distributions of the natural frequency and damping ratio are found to be less sensitive to the data length and the number of data sets, compared to the distribution of the PSD of modal force and PSD of prediction error. One possible reason is that the latter two parameters are more sensitive to environmental conditions.

Based on the probabilistic model, the distribution of the modal properties in a future time window under environment covered by one of the time segments obtained in the monitoring database can be assessed. This makes it possible to estimate the variability of the modal parameters in a similar operational environment in the future and then predict the dynamic characteristics of subject structure utilizing the data collected. The mean and variance of the prediction model are also derived based on the distribution. The c.o.v. values obtained tend to be larger than the posterior c.o.v. calculated utilizing single data set. This is reasonable since the prediction model takes into account the variability among different data sets. The data length and the number of data sets are two important quantities. The choice is a balance between modeling error (e.g., stationarity) and identification precision. It will be an interesting topic to investigate how much data is adequate to establish a reliable prediction model.

Acknowledgements

The work in this paper is partially supported by Grant EGG10034 from the University of Liverpool, the National Basic Research Program of China (973 Program) (Project No. 2014CB049100) and Fundamental Research Funds for the Central Universities, China (Grant No. 2014KJ040).

Appendix: Derivation of expectation and standard derivation

Let $E_D[\cdot]$ denote the expectation when Θ is distributed as $p_\Theta(\theta \mid D)$ given by equation (9). The mean of Θ can be derived as follows.

$$\begin{aligned}
E_D[\Theta] &= \int_{-\infty}^{\infty} \theta \; p_\Theta(\theta \mid D) d\theta \\
&= \int_{-\infty}^{\infty} \theta \; \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{1}{\sqrt{2\pi \hat{c}_i}} \exp\left[-\frac{1}{2\hat{c}_i} (\theta - \hat{\theta}_i)^2 \right] d\theta \\
&= \frac{1}{N_s} \sum_{i=1}^{N_s} \int_{-\infty}^{\infty} \theta \; \frac{1}{\sqrt{2\pi \hat{c}_i}} \exp\left[-\frac{1}{2\hat{c}_i} (\theta - \hat{\theta}_i)^2 \right] d\theta \\
&= \frac{1}{N_s} \sum_{i=1}^{N_s} \hat{\theta}_i \\
&= \text{sample mean of } \{\hat{\theta}_i\}_{i=1}^{N_s}
\end{aligned} \tag{29}$$

Let $\mathrm{var}_D[\Theta]$ denote the variance of Θ when it is distributed as $p_\Theta(\theta \mid D)$. Also, let Θ_i denote a Gaussian random variable with mean $\hat{\theta}_i$ and variance $\hat{c}_i$. Then

$$\begin{aligned}
&\mathrm{var}_D[\Theta] = E_D[\Theta^2] - E_D[\Theta]^2 \\
&= \frac{1}{N_s} \sum_{i=1}^{N_s} E[\Theta_i^2 \mid D] - E_D[\Theta]^2 \\
&= \frac{1}{N_s} \sum_{i=1}^{N_s} [\hat{\theta}_i^2 + \hat{c}_i] - E_D[\Theta]^2 \\
&= \frac{1}{N_s} \sum_{i=1}^{N_s} \hat{c}_i + \frac{1}{N_s} \sum_{i=1}^{N_s} \hat{\theta}_i^2 - \left(\frac{1}{N_s} \sum_{i=1}^{N_s} \hat{\theta}_i\right)^2 \\
&= \text{sample mean of } \{\hat{c}_i\}_{i=1}^{N_s} + \text{sample variance of } \{\hat{\theta}_i\}_{i=1}^{N_s}
\end{aligned} \tag{30}$$

References

Au S.K. (2011). “Fast Bayesian FFT method for ambient modal identification with separated modes.” Journal of Engineering Mechanics, 137(3), 214-226.

Au S.K. (2012a). “Fast Bayesian ambient modal identification in the frequency domain, Part I: Posterior most probable value.” Mechanical Systems and Signal Processing, 26(1), 60-75.

Au S.K. (2012b). “Fast Bayesian ambient modal identification in the frequency domain, Part II: Posterior uncertainty.” Mechanical Systems and Signal Processing, 26(1),76-90.

Au S.K. (2012c). “Connecting Bayesian and frequentist quantification of parameter uncertainty in system Identification. ” Mechanical Systems and Signal Processing 29, 328-342.

Au S.K., Ni Y.C., Zhang F.L. and Lam H.F. (2012a). “Full-scale dynamic testing and modal identification of a coupled floor slab system. ” Engineering Structures, 37, 167-178

Au S.K. and To P. (2012). “Full-Scale validation of dynamic wind load on a super-tall building under strong wind. ” Journal of Structural Engineering,138(9), 1161-1172

Au S.K., Zhang F.L. (2012a). “Fast Bayesian ambient modal identification incorporating multiple setups. ” Journal of Engineering Mechanics, ASCE,138(7), 800–815.

Au S.K. and Zhang F.L. (2012b). “Ambient modal identification of a primary-secondary structure using fast Bayesian FFT approach. ” Mechanical Systems and Signal Processing, 28, 280-296.

Au S.K., Zhang F.L., Ni Y.C. (2013). “Bayesian operational modal analysis: theory, computation, practice. ” Computers and Structures, 126, 3-14.

Au S.K., Zhang F.L. and To P. (2012). “Field observations on modal properties of two tall buildings under strong wind. ” Journal of Wind Engineering and Industrial Aerodynamics,101, 12-23.

Bashor R and Kareem A (2008). "Uncertainty in damping and damping estimates: An assessment of database and data from recent full-scale measurements." 2008 Structures Congress, Structural Engineering Institute (SEI), Reston, VA.

Beck J.L. (2010). “Bayesian system identification based on probability logic. ” Structural Control and Health Monitoring 17 (7), 825-847

Bentz A. and Kijewski-Correa T. (2008). "Predictive models for damping in buildings: The role of structural system characteristics." 2008 Structures Congress, 18th Analysis and

Computation Specialty Conference, Structural Engineering Institute (SEI), Reston, VA

Brewick P.T., Smyth A.W. (2014). “On the application of blind source separation for damping estimation of bridges under traffic loading”, Journal of Sound and Vibration, 333(26): 7333-7351.

Brownjohn J.M.W. (2003). “Ambient vibration studies for system identification of tall buildings. ” Earthquake Engineering and Structural Dynamics, 32, 71-95.

Brownjohn J.M.W., Magalhaes F., Caetano E., Cunha A. (2010). “Ambient vibration re-testing and operational modal analysis of the Humber Bridge. ” Engineering Structures, 32, 2003-2018.

Catbas F.N. et al. (2011). “Structural Identification (St-Id) of constructed facilities. ” ASCE SEI. www.cece.ucf.edu/people/catbas/St ID Report.

Çelebi M. (1996). "Comparison of damping in buildings under low-amplitude and strong motions." Journal of Wind Engineering and Industry Aerodynamics, 59 (2-3): 309-323.

Clough R.W., Penzien J. (1993). *Dynamics of Structures*, 2nd ed. McGraw-Hill, New York.

Davenport A.G., Hill-Carroll P. (1986). “Damping in tall buildings: its variability and treatment in design. ” In Proceedings of ASCE Spring Convention, Seattle, USA, Building Motion in Wind, pp.42–57.

Dohler M., Lam X.-B., Mevel L. (2013). “Uncertainty quantification for modal parameters from stochastic subspace identification on multi-setup measurements. ” Mechanical Systems and Signal Processing, 36, 562–581.

El-kafafy M., De Troyer T., Peeters B., Guillaume P. (2013). “Fast maximum-likelihood identification of modal parameters with uncertainty intervals: A modal model-based formulation. ” Mechanical Systems and Signal Processing, 37, 422–439.

Fang J.Q., Li Q.S., Jeary A.P. and Liu D.K. (1999). "Damping of tall buildings: Its evaluation and probabilistic characteristics." The Structural Design of Tall Buildings, 8 (2), 145-153.

Jeary A.P. (1997). “Damping in structures. ” Wind Engineering and Industrial Aero-dynamics 72,345–355.

Kijewski-Correa T., et al. (2006). “Validating wind-induced response of tall buildings: Synopsis of the Chicago full-scale monitoring program. ” Journal of Structural Engineering, 132(10), 1509–1523.

Liu Y., Duan Z.D. (2002). “Fuzzy finite element model updating of bridges by considering the uncertainty of the measured modal parameters. ” Science China Technological Sciences, 55 (11), 3109–3117.

Ni Y.Q., Hua X.G., Fan K.Q., Ko J.M. (2005). “Correlating modal properties with temperature using long-term monitoring data and support vector machine techniques. ” Engineering Structures, 27(12), 1762-1773.

Nishio M., Marin J., Fujino Y. (2012). “Uncertainty quantification of the finite element model of existing bridges for dynamic analysis. ” Journal of Civil Structural Health Monitoring, 2 (3-4), 163-173.

Papadimitriou C., Beck J. L., and Katafygiotis L. S. (2001). “Updating robust reliability using structural test data. ” Probability Engineering Mechanics, 16(2), 103–113.

Pintelon R., Guillaume P., Schoukens J. (2007). “Uncertainty calculation in (operational) modal analysis. ” Mechanical Systems and Signal Processing, 21, 2359–2373.

Ramos L.F., Marques L., Lourenco P.B., De Roeck G., Campos-Costa A., Roque J. (2010). “Monitoring historical masonry structures with operational modal analysis: Two case studies. ” Mechanical Systems and Signal Processing 24, 1291–1305.

Reynders E., Pintelon R., De Roeck G. (2008). “Uncertainty bounds on modal parameters obtained from stochastic subspace identification. ” Mechanical Systems and Signal Processing, 22, 948–969.

Reynders E. (2012). “System Identification Methods for (Operational) Modal Analysis: Review and Comparison. ” Arch Comput Methods Eng, 19, 51–124.

Satake N., Suda K., Arakawa T., Sasaki A., Tamura A. (2003). “Damping evaluation using full-scale data of buildings in Japan. ” Journal of Structural Engineering 129, 470–477.

Schoukens J, Pintelon R. (1991). *Identification of linear systems: a practical guideline for accurate modelling.* London: Pergamon Press.

Shi W., Shan J., Lu X. (2012). “Modal identification of Shanghai World Financial Center both from free and ambient vibration response. ” Engineering Structures, 36: 14–26.

Spence S. and Kareem A. (2014). "Tall Buildings and Damping: A Concept-Based Data-Driven Model." Journal of Structural Engineering, 140(5), 04014005.

Steenackers G., Guillaume P. (2006). “Finite element model updating taking into account the uncertainty on the modal parameters estimates. ” Journal of Sound and Vibration, 296, 919–934.

Wenzel H., Pichler D. (2005). *Ambient vibration monitoring*, John Wiley & Sons, UK.

Whelan M.J., Gangone M.V., Janoyan K.D., Jha R. (2009). “Real-time wireless vibration monitoring for operational modal analysis of an integral abutment highway bridge. ” Engineering Structures, 31, 2224-2235.

Xia Y., Hao H., Zanardo G., Deeks A. (2006). “Long term vibration monitoring of an RC slab: temperature and humidity effect. ” Engineering Structures, 28(3), 441-452.

Yang Y. and Nagarajaiah S. (2013). "Time-Frequency Blind Source Separation Using Independent Component Analysis for Output-Only Modal Identification of Highly Damped Structures." Journal of Structural Engineering, 139, SPECIAL ISSUE: Real-World Applications for Structural Identification and Health Monitoring Methodologies, 1780-1793.

Yuen K.V., Katafygiotis L.S. (2003). “Bayesian Fast Fourier Transform approach for modal updating using ambient data. ” Advances in Structural Engineering, 6(2), 81-95.

Yuen K.-V., Kuok S.-C. (2010). “Ambient interference in long-term monitoring of buildings. ” Engineering Structures, 32, 2379-2386.

Zhang F.L. and Au S.K. (2013). “Erratum to 'Fast Bayesian FFT method for ambient modal identification with separated modes' by Siu-Kui Au.” Journal of Engineering Mechanics,139(4), 545-545.

Zhang F.L., Au S.K. and Lam H.F. (2014). “Assessing uncertainty in operational modal analysis incorporating multiple setups using a Bayesian approach.” Structural Control and Health Monitoring. DOI: 10.1002/stc.1679. In print.

Table 1 Sample and Bayesian statistics, SDOF synthetic data

	Exact	Single Set MPV	Sample mean of MPV	Freq.[a] (%) A	Bay.[b] (%)B	A/B
f (Hz)	1	1.001	1.000	0.235	0.239	0.99
ζ (%)	1	0.875	1.063	25.15	25.72	0.98
$S((\mu g)^2/Hz)$	1	0.980	1.018	19.61	23.62	0.83
$S_e((\mu g)^2/Hz)$	100	103.996	100.293	5.16	5.07	1.02

[a]Frequentist=sample c.o.v. of MPV= (sample std. of MPV)/ (sample mean of MPV).
[b]Bayesian= (r.m.s. of posterior std.)/ (sample mean of MPV).

Table 2 C.o.v.s calculated based on single data set in four different cases, SDOF synthetic data

Case	f (%)	ζ (%)	S (%)	Se (%)
a)	0.20	24.6	20.3	4.9
b)	0.26	23.4	21.9	5.1
c)	0.19	17.9	17.9	3.7
d)	0.33	40.4	40.7	7.2

Table 3 Sample and Bayesian statistics, lab shear frame data

	Single Set MPV	Sample mean of MPV	Freq.[a] (%) A	Bay.[a] (%)B	A/B
f (Hz)	1.379	1.380	0.058	0.070	0.83
ζ (%)	0.181	0.236	27.19	29.99	0.91
$S((\mu g)^2/Hz)$	4.01×10^3	3.40×10^3	17.46	6.32	2.76

[a]See Table 1

Table 4 Sample and Bayesian statistics, super tall building

	Single Set MPV	Sample mean of MPV	Freq.[a] (%) A	Bay.[a] (%)B	A/B
f (Hz)	0.154	0.154	0.20	0.17	1.18
ζ (%)	0.615	0.499	33.07	32.30	1.02
$S((\mu g)^2/Hz)$	0.744	0.210	98.31	13.15	7.47

[a]See Table 1

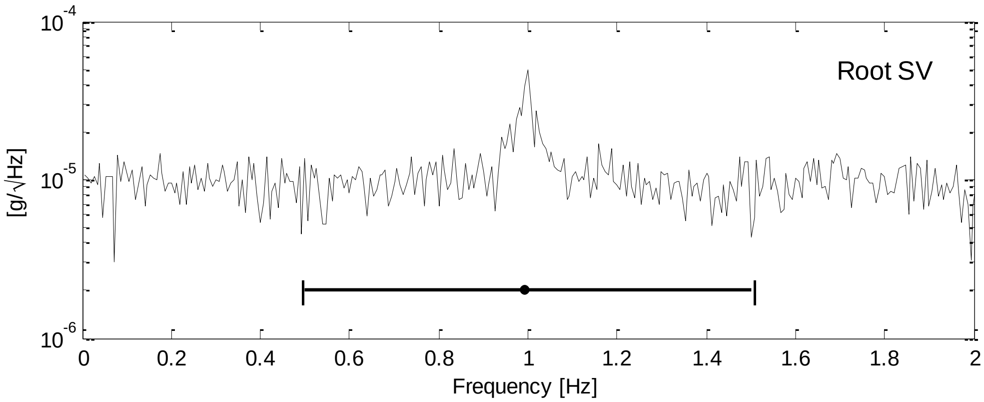


Figure 1 Root singular value (SV) spectrum of a typical data set, synthetic data

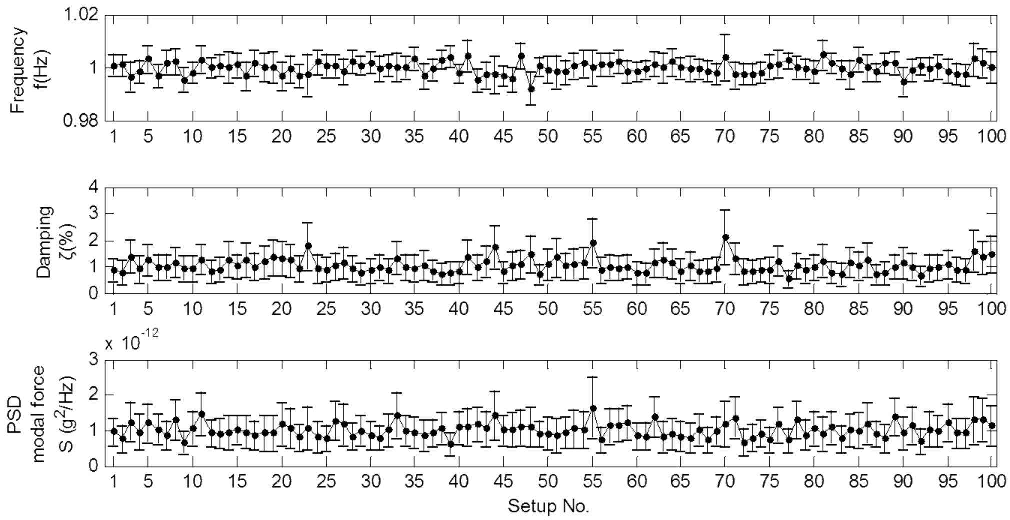
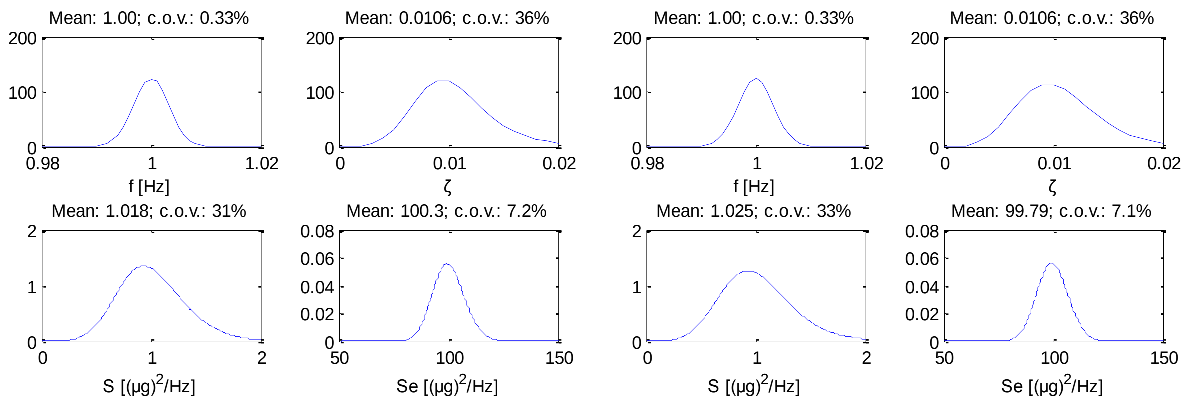
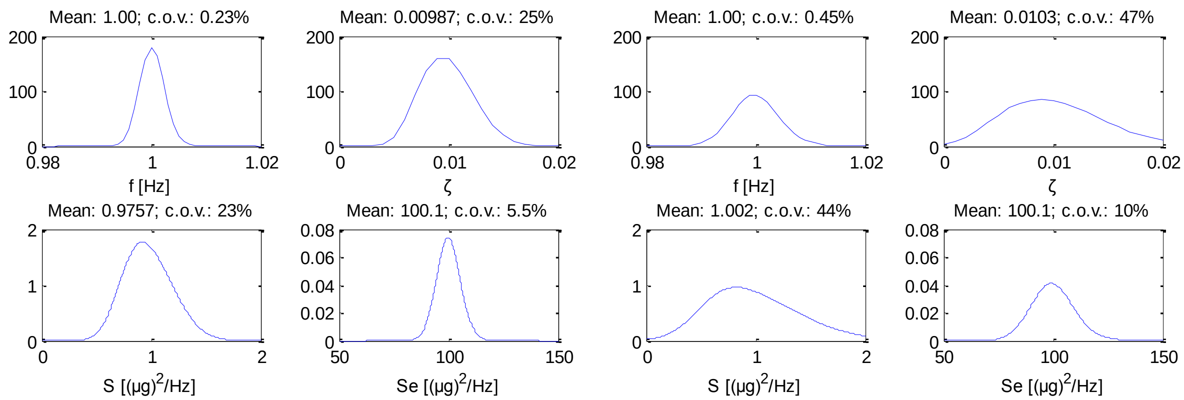


Figure 2 Modal identification results of a SDOF structure in different setups, Synthetic data



a).100 sets data, 600sec. b). 1000 sets data, 600sec.



c). 100 sets data, 1200sec d). 100 sets data, 300sec.

Figure 3 PDF of the modal parameters based on the proposed probabilistic model

Figure 4. Shear frame in the laboratory

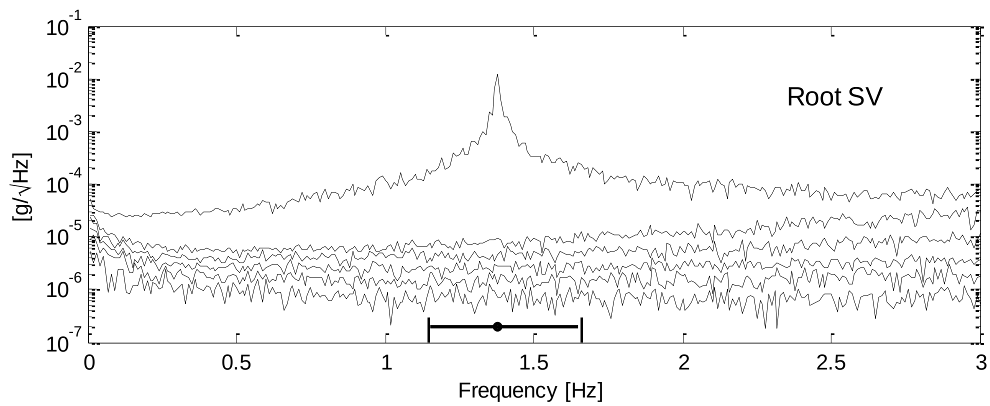


Figure 5 Root singular value (SV) spectrum of a typical data set, laboratory data

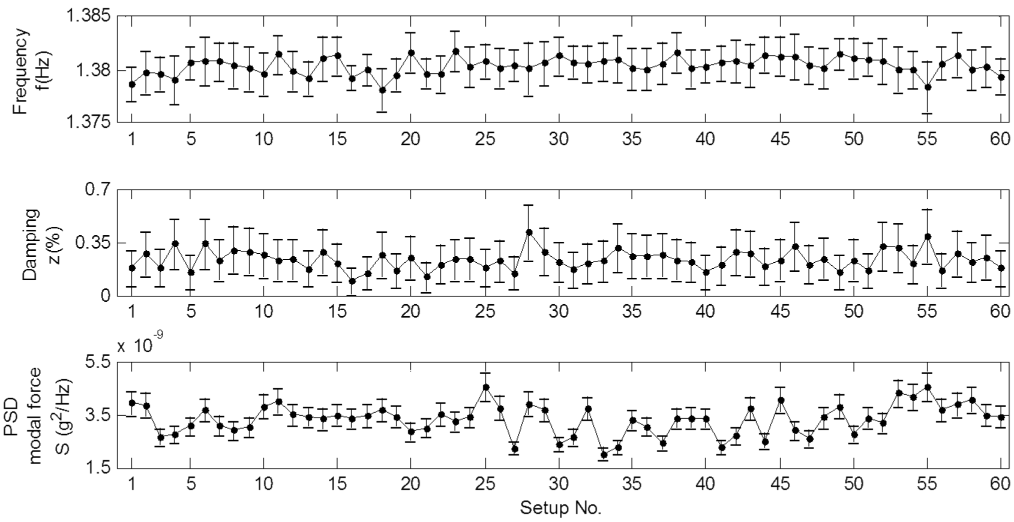
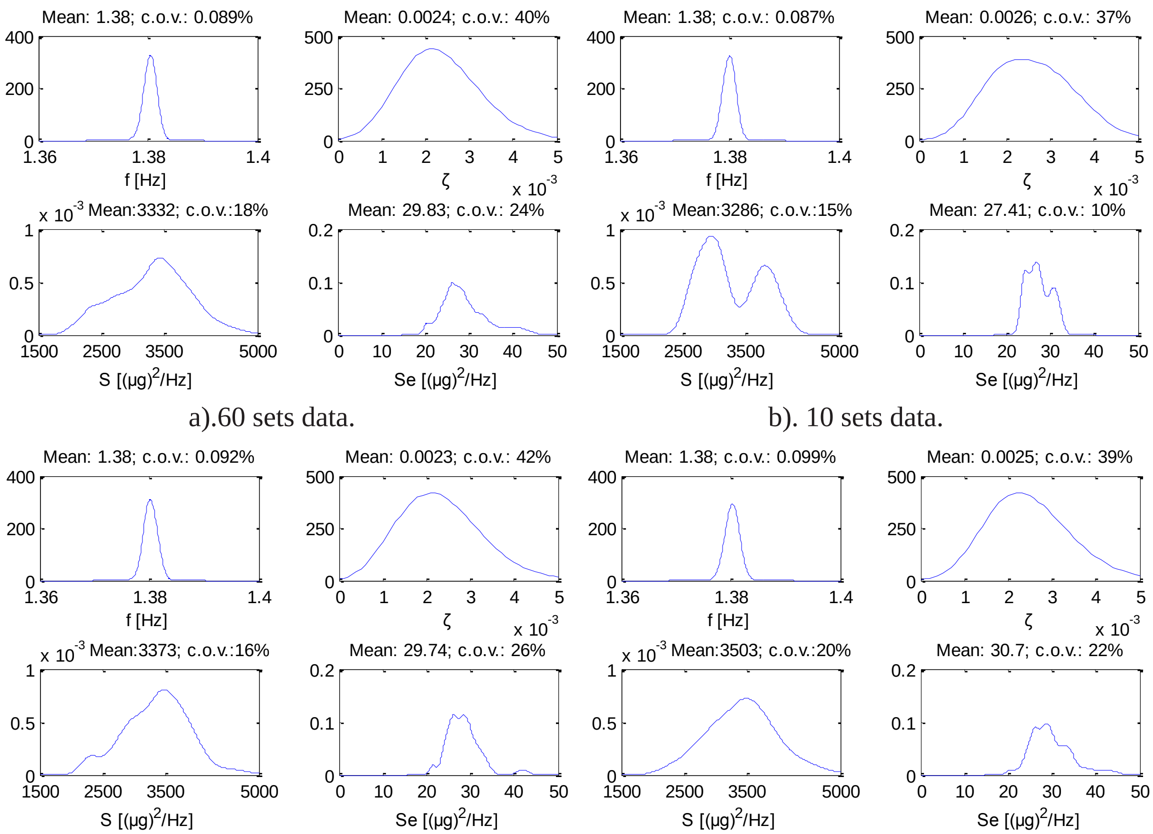


Figure 6. Modal identification results of a lab frame in different setups arranged chronologically, data from four uni-axial accelerometer



a).60 sets data. b). 10 sets data.

c). 30 sets data d). 120 sets data.

Figure 7 PDF of the modal parameters based on the proposed probabilistic model

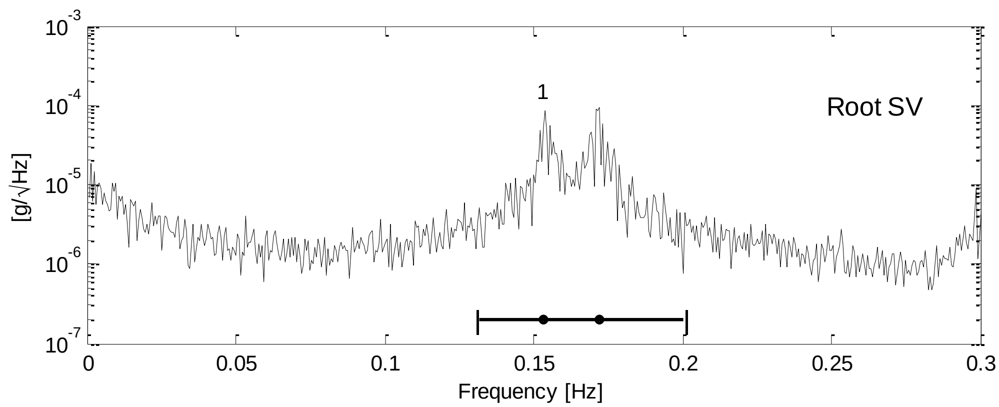


Figure 8 Root singular value (SV) spectrum of a typical data set, super tall building

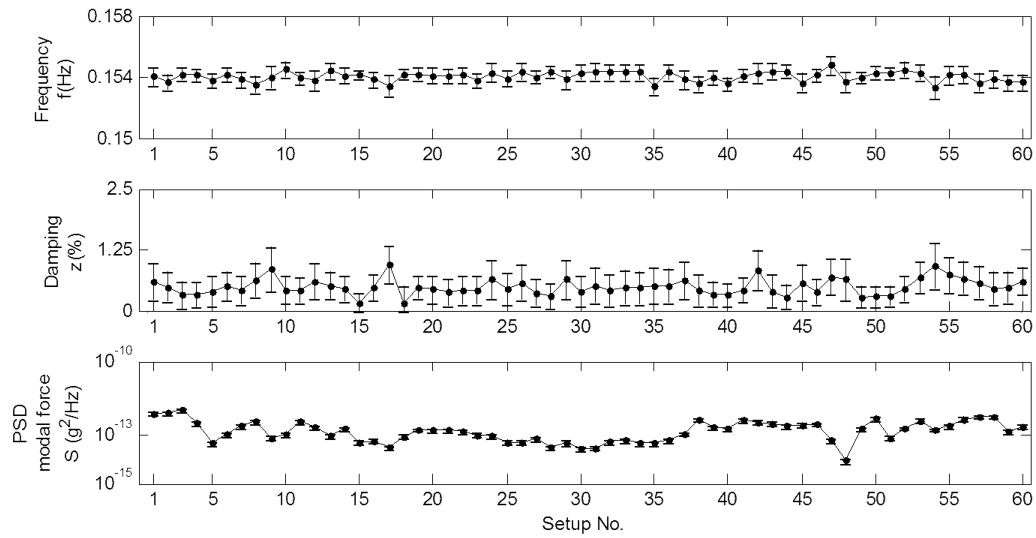


Figure 9. Modal identification results of a super tall building in different setups arranged chronologically, data from one tri-axial accelerometer

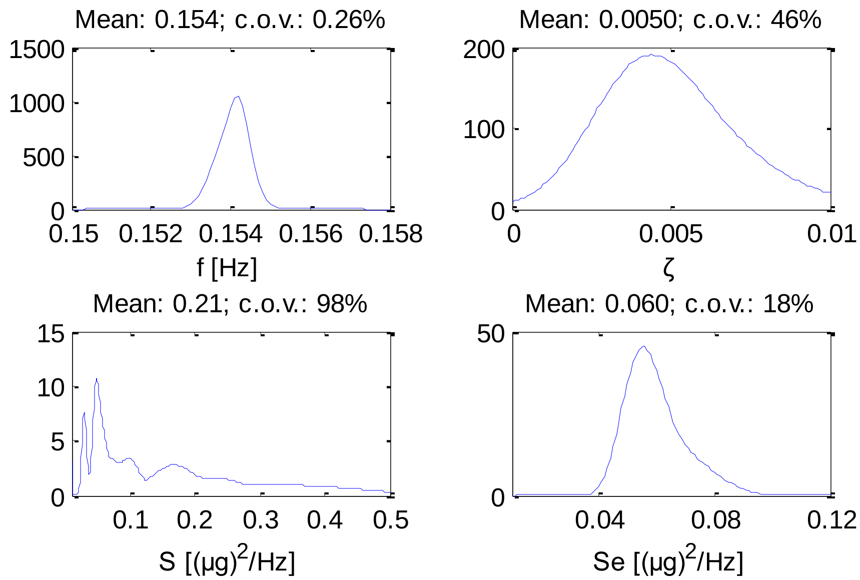


Figure 10 PDF of the modal parameters based on the proposed probabilistic model, super tall building.